

METHOD AND APPARATUS FOR CLASSIFYING JAVA REMOTE METHOD INVOCATION TRANSPORT TRAFFIC

INVENTORS

5

Robert J. Duncan, Tal I. Lavian

FIELD OF THE INVENTION

This invention generally relates to exchanging information over a data communications network. More particularly, the invention relates to the preferential exchange of 10 information including remote method invocations over a network implementing a differentiated service mechanism. Even more particularly, the invention relates to such preferential remote method invocations using the Java programming language.

BACKGROUND OF THE INVENTION

Remote procedure calls (RPCs) are widely used in data communications, particularly in layer 5 of OSI architecture-type systems. An RPC is a request conforming to a protocol that a program can use to request a service from a program on another computer on the network. The client program making the request includes an API which receives calls to remote procedures from other parts of the program and passes them to a runtime program in the client. 20 The runtime program in the client knows the format for addressing the remote server and relays the procedure request accordingly.

Generally speaking, remote method invocations (RMIs) are RPCs which are written in the Java programming language. Unlike their RPC counterparts, RMIs are object-oriented and can pass objects along with the RMI requests, as well as receive objects as results.

Thus, RMIs are more appropriate for distributed object-oriented systems where communication between objects at different addresses on one or more networks is needed. In a typical RMI application, a server application creates a number of remote objects (remote to the client systems), makes references to those objects accessible via a registry, and allows clients to invoke

- 5 methods on those remote objects. An application on a client system gets references to these remote server objects and invokes their methods. In this scheme, RMI is the way in which the client and server communicate with respect to these invoked methods. Like RPC, these layers are generally located between the transport and application layers.

Differentiated service (DiffServ) enabled networks have become more popular in recent years. DiffServ is a network service mechanism that classifies network traffic into a number of differentiated service classes and controls the network traffic so that certain classes get precedence or service quality. The preference relationships are defined by service level agreements (SLAs) among and between network applications and service providers. For example, one SLA can provide that certain classes of traffic between two networks, such as important business applications, should be prioritized over other classes of traffic, such as web surfing, between the networks. As another example, an SLA can specify that a certain class of traffic from a first network to a second network should be prioritized over a certain class of traffic from the first network to a third network. DiffServ-enabled devices in a network use six-bit Differentiated Services Code Point (DSCP) labels included in IP packet headers to know how 20 to forward the packet.

Related to the concept of DiffServ is that of edge routers. An edge router is a device for forwarding data between two networks to which it is connected. Edge routers are

used to classify data flows of a DiffServ network by comparing packets in such flows with relevant SLA information stored in the router. The edge router receives packets bound for a DiffServ network and a quality of service (QoS) classifier within the edge router and applies DSCP labels according to the SLA specification. If a packet belongs to a flow with a high

- 5 priority, it receives a DSCP corresponding to that high priority; if the packet belongs to a flow with a low priority, it receives a DSCP corresponding to that low priority.

The packet is then delivered to an external network for eventual transmission to its ultimate destination according to the DSCP requirements. If the DiffServ network becomes congested, the packets labeled with high priority DCSPs are more likely to successfully reach their destinations; packets labeled with low priority DCSPs might only receive best effort service. This helps to guarantee a better quality of service for higher priority flows.

RMI calls may be hard to detect by QoS classifiers in edge routers; thus, an RMI which should be given a high priority DSCP might not be properly classified, resulting in a failure in QoS requirements. This is because QoS classifiers usually screen packets in accordance with a five-tuple that includes a socket port number. RMI flows generally do not always use well-known port numbers. This creates a problem because priority levels which should be accorded to RMIs may then be usurped by other applications that can more easily be screened for DiffServ classification and transmission.

20

SUMMARY OF THE INVENTION

A data communication network for DiffServ communications in which a customized Java socket factory is added to clients connected to a data communication network

having a DiffServ-enabled edge router. When an application running on a client system wishes to make a remote procedure call to a remote server system on another network, it makes its usual call to an RMI stub which invokes an RMI transport layer having a custom socket factory to generate a socket used in the RMI call. The custom socket factory detects when a high priority

- 5 RMI call is being made and can determine the identity of the calling procedure as well. The socket factory makes a side channel communication to the edge router, or an alternated network resource broker and a service policy maker, to provide this RMI call information to the edge router, which then makes use of this data when performing DiffServ classification for packets transmitted during the course of the call.

10
15
20

BRIEF DESCRIPTION OF THE DRAWINGS

These and other aspects and features of the present invention will become apparent to those ordinarily skilled in the art upon review of the following description of specific embodiments of the invention in conjunction with the accompanying figures, wherein:

FIGURE 1 shows structure of a network according to an embodiment of the present invention;

FIGURE 2 shows structure of a client system according to an embodiment of the present invention;

FIGURE 3 shows operation of an RMI transport layer according to an embodiment of the invention;

FIGURE 4 shows structure in an edge router according to an embodiment of the present invention; and

FIGURE 5 shows operation of an embodiment of the present invention.

DETAILED DESCRIPTION OF EMBODIMENTS OF THE INVENTION

The present invention will now be described in detail with reference to the

5 drawings, which are provided as illustrative examples of the invention so as to enable those skilled in the art to practice the invention. Notably, the implementation of certain elements of the present invention may be accomplished using software, hardware or any combination thereof, as would be apparent to those of ordinary skill in the art, and the figures and examples below are not meant to limit the scope of the present invention. Moreover, where certain elements of the present invention can be partially or fully implemented using known components, only those portions of such known components that are necessary for an understanding of the present invention will be described, and detailed descriptions of other portions of such known components will be omitted so as not to obscure the invention. Further, the present invention encompasses present and future known equivalents to the known components referred to herein 10 by way of illustration.

15

FIG. 1 illustrates the architecture of one embodiment of the present invention. As shown in FIG. 1, the architecture includes a client system 102 that has been adapted to communicate with a remote server 104 via an edge router 106 and inside and outside networks 110 and 108. Such communications between the client system 102 and the remote server 104 20 can include, for example, RMIs as described above.

The client system 102 can be, for example, a workstation such as a PC, Macintosh, Unix, Linux or Sparc workstation. The server 104 can be, for example, a PC,

Macintosh, Unix, Linux or Sparc workstation, mainframe, minicomputer like the IBM AS-400, or any other appropriate server hardware. The software on the server can be any appropriate server application, or any application running on an application server like IBM's San Francisco, Sun's Net Dynamics, Bluestone, BEA or any appropriate application server, Database servers

5 that can run on these hardware servers include Oracle, IBM's DB2, an object repository (server) or other appropriate system. The inside network 110 can be, for example, a LAN or other network connecting the client system 102 and router 106. The edge router 106 can be, for example, a Nortel Networks Passport 8600, Cisco System Catalyst 6500 or other router which includes DiffServ or other network control functionality like RSVP, COPS, COPS-PR, MPLS, IntServ or similar functionality. The outside network 108 can be a LAN, MAN, WAN or the Internet, for example. It further preferably includes one or more DiffServ-enabled devices. Specifically, at least part of the path between the client 102 and the remote server 104 is managed using differentiated services.

Preferably, the router 106 is at the edge of the inside network 110, but this is not necessary. What is desirable is that this is the first or closest DiffServ-classifying router in the DiffServ-controlled portion of the path between the client 102 and the server 104. Typically, this will be at the edge of the inside network 110. The invention may be practiced in a router within the outside network 108; however, its benefit (as well as the benefit of any DiffServ-capable QoS) will be less there because part of the transit path of RMI traffic will be non-DiffServ prioritized.

The present invention is preferably practiced instead of the aforementioned conventional DiffServ capabilities known in the art. Alternatively, the edge router 106

preferably has existing DiffServ-classifying capability in accordance with SLAs or it has been upgraded with dynamic DiffServ-classifying capability, and the present invention is practiced in addition to, these conventional DiffServ classifying capabilities in accordance with SLAs.

Generally, when operating in accordance with the invention, a client application

- 5 in the client system 102 initiates an RMI connection with the remote server 104. The present invention, preferably implemented as software within the client system 102, detects the attempt to initiate an RMI connection using special software described below (alternatively, the client and server can decide that all RMI connection traffic will be on an arbitrary TCP port) and collects RMI flow information relating to the connection (e.g., a five-tuple including the sender and receiver MAC and IP addresses, the sender and receiver MAC and IP port numbers, and the TCP protocol type for the connection) for transmission via a side channel to the edge router 106. Although other information sets may be usable, this arrangement allows the system to distinguish specific traffic. This is part of the interaction between the application and the routers 106. The routers 106 thus have a specific way to identify specific traffic and give the right service to it.
- 10 15

- A DiffServ classifier 408 (see FIG. 4) in the edge router 106 becomes aware of the received flow information and assigns the RMI flow a QoS level in accordance with the SLAs. Thereafter, when communications between the client system 102 and the remote server 104 associated with the RMI flow are received by the edge router 106, they are marked with the appropriate DiffServ labels for the desired QoS by the DiffServ edge router 106 and are forwarded through the outside network 108. For example, the old ToS byte may be used as a DSField byte for marking the packets.
- 20

Although only one client system 102 and edge router 106 are shown, it should be noted that the remote server 104 preferably includes functionality for communicating with many client systems 102 via the same or many different edge routers 106. These may be a mix of devices practicing the invention, DiffServ-capable devices not practicing the invention, and non-

- 5 DiffServ capable devices. Likewise, the client system 102 can include functionality for communicating with many different edge routers 106 and servers 104. Such additional functionality is apparent to those skilled in the art and more detailed descriptions thereof are not necessary for an understanding of the present invention.

It should be further noted that the remote server 104 and its associated edge router or similar device (not shown) may include similar functionality as described herein with reference to the client system 102 and the edge router 106 for dynamically assigning QoS to RMI return flows from the server 104 to the client system 102. Alternatively, the client system 102 may include functionality for also informing the remote server's associated edge router of the RMI flow information for dynamic QoS assignment.

10 It should be even further noted that, although the present invention is described herein with reference to an exemplary implementation in a DiffServ-capable network, the present invention is not limited to this example. Rather, the present invention is applicable to other existing or emerging standards for providing differentiated services for different types of flows or classes of network traffic, such as IntServ/RSVP, COPS, COPS-PR, MPLS, IEEE 801.2(p) 15 and IEEE 801.2(q). Moreover, where an outside network 108 supports more than one type of differentiated services, the client system 102 may include functionality for dynamically adjusting the level of service for RMI flows in accordance with these multiple types.

FIG. 2 illustrates an example client system 102 in accordance with one embodiment of the present invention. As shown in this example implementation, the client system 102 includes software modules 202 to 212 that can be provided as one or more executable files loaded in a dynamic memory 220 and executed by a central processing unit 222 in the client system 102. The client system 102 further includes a user interface 224 that allows an associated user such as a network manager to initiate an RMI connection and a network interface 226 through which the client system 102 can communicate via the inside network 110 with the edge router 106 and the outside network 108, as well as a bus 228 that allows communications between the memory 220, the CPU 222, the user interface 224 and the network interface 226.

As shown in FIG. 2, the software loaded in the memory 220 during operation of the client system 102 includes a client application 202 that establishes connections with remote objects on the remote server 104 for the purposes of making RMIs, for example. As indicated in the Figure, the client application 202 may communicate with an application library 204 either directly or through one or more additional software layers (not shown).

The application library 204 is typically one or more APIs that are designed to communicate with an RMI transport layer 208 so the client application need not be aware of the implementation details of the RMI transport. The library 204 can include, for example, Java programs and the like, that provide an API and interface functionality and implement or initiate RMIs on the remote server 104. As is known in the art, the library 204 may be provided by the same entity that maintains the remote server 104 to enable the client system 102 to establish and maintain connections with that server.

The RMI transport layer 208 further communicates with a socket / TCP layer 210 which may even further communicate with other networking protocol layers and device-level layers such as a device operating system and device-level APIs. These layers are device-specific and are generally specifically written for the device and operating system 212 (e.g., Solaris, 5 Windows) of the client system 102. Further detailed descriptions thereof are not necessary for an understanding of the present invention.

In accordance with the invention, as shown in FIG. 2, the software load further includes an RMI stub 206. Generally speaking, the stub 206 receives calls from other Java program code directing the RMI transport layer 208 to make RMI calls and passes the traffic on 10 to the RMI transport layer 208 for transport processing. The primary function of the RMI stub is to act as a "straw man" that looks to calling programs in the application library 204 like an interface of the remote objects being called. In this way, the remote access functionality can be kept largely transparent to system applications.

It should be noted that the client system 102 can include additional software to 15 that shown in FIG. 2, and that certain of the software layers can include functionality for communicating with other software layers not shown in FIG. 2, as well as additional functionality and other software not necessary for practicing the present invention.

It should be further noted that the RMI stub 206 can be integrated together with, or (for exclusive DiffServ operation) entirely replace, the RMI transport layer 208. Many other 20 combinations or divisions between the functionalities of the RMI stub 206 and other software layers are possible, as will become apparent to those skilled in the art after being taught by the present specification.

FIG. 3 is a block diagram illustrating an example of functionality within the RMI transport layer 208 in accordance with one embodiment of the present invention. Before examining the operation of the RMI transport layer 208 in detail, it will be helpful to briefly describe the server-side process of preparing remotely accessible objects for use by client systems 102.

5 Referring to FIG. 3, in a typical RMI application a server 104 application creates, has loaded, etc. a number of remote objects (remote to the client systems 102) 302 and calls a registry database 304 to provide Java references for the objects 302 thereto. The registry 304 associates each object reference with a name that will be known by a client system 102 and stores the association therein. Then, when making an RMI call, the client system 102 accesses the registry 304 using an object's name to get its reference on the server 104. Using the object reference, the client system 102 can invoke methods of the referenced object 302 on the remote server 104.

Referring again to the RMI transport layer 208, when the application library 204 calls the stub 206 to make an RMI call, the RMI stub 206 invokes the RMI transport layer 208 which calls a socket factory 306 therein to provide a socket for use in establishing the communication session. A conventional socket factory would generate a socket according to predetermined rules of the client system 102, edge router 106 and network and provide it to the RMI transport layer 208 for use in the communication session. A socket factory 306 according to an embodiment of the present invention, however, checks to see whether the requested connection is for an RMI request, e.g., to a predefined socket number, and if so additionally passes the socket number and other flow information, e.g., the five-tuple of the client system 102, (the client or the server can initiate this information) to the edge router 106 via a side

channel, e.g., a Java servlet, and identifies the information as pertaining to high priority RMI traffic. The edge router 106 then makes use of this flow information in DiffServ operations as described in greater detail below. The custom socket factory 306 further passes the socket and other relevant information to the RMI transport layer 208 to initiate the RMI call.

5 It should be noted that as seen above, the presence and operation of this embodiment may be transparent to the application library 204 -- it simply makes the API calls and does not know (or care) that they are passed to the edge router 106.

FIG. 4 illustrates an example of the RMI QoS functionality in the edge router 106 in accordance with one embodiment of the present invention. As shown, it includes an RMI signal interface 402, an RMI handler 404, a dynamic RMI list 406, a DiffServ classifier 408 and a service level agreement (SLA) store 410. Although shown separately for clarity, it should be appreciated that the functionalities of certain elements can be combined together and divided apart in various ways.

The RMI signal interface 402 communicates with the RMI transport layer 208 in the client system 102. For example, the RMI signal interface 402 can be the portion of a CGI script (written in PERL, C/C++, etc.) that executes when the RMI transport layer 208 posts information corresponding to the script to the edge router 106. The RMI signal interface 402 may extract the flow information from the CGI script parameters. It should be apparent that the edge router 106 may include additional functionality for communicating with the RMI transport layer 208 using known protocols such as HTTP, for example, and for launching CGI applications in response to information posted to the edge router 106.

The RMI handler 404 receives the RMI flow information from the RMI signal interface 402 and adds it to a dynamic RMI list 406. The dynamic RMI list 406 contains entries listing RMI flows that have been requested through the edge router 106. In one example of the invention, the RMI handler 404 includes functionality for aging out entries in the list 406 using a 5 least-recently-used (LRU) algorithm or the like. Alternatively, the RMI handler 404 simply overwrites the oldest entry when the list becomes full and new flow information is received.

The SLA store 410 contains QoS settings among and between networks in accordance with service level agreements (SLAs). These are the static and pre-defined flows for conventional DiffServ treatment. In one example of the invention, the dynamic RMI list 406 and the SLA store 410 are provided as separate memories, and are shown as such for clarity. However, this is not necessary. They can be provided together in one memory area, with a static portion (locations not addressed during operation by the RMI handler 404, corresponding to the SLAs in SLA store 410) and a dynamic portion (locations addressable during operation by the RMI handler 404 to update with RMI flow information, corresponding to the dynamic RMI list 15 406).

The DiffServ classifier 408 accesses the flow information in the dynamic RMI list 406 and the SLA store 410 and compares the information therein with flow information detected from packet headers of traffic flowing to the outside network 108 through the edge router 106. If the flow information in the traffic headers matches the stored flow information (e.g., the five- 20 tuple of source and destination IP addresses, source and destination MAC addresses, and socket or port number), the packets are marked with a QoS label in accordance with the QoS agreed to for that flow. The classifier 408 can be implemented by adapting a conventional DiffServ

classifier with the additional functionality of comparing packet headers with dynamic RMI flow information stored in the dynamic RMI list 406. It should be apparent, however, that such adaptation can be minimal or unnecessary if the flow information in the dynamic RMI list 406 and the SLA store 410 are provided together in a single contiguous memory location.

5 It should be noted that the edge router 106 preferably further includes functionality for authenticating requests for dynamic classification of RMI flows. This functionality can include encryption and signature techniques. Such additional functionality that presently exists is well understood by those skilled in the art and the details thereof are not necessary for an understanding of the present invention.

10 Moreover, the edge router 106 can also include a conventional packet processing and forwarding architecture, a further description of which is not necessary for an understanding of the present invention.

FIG. 5 is a flowchart illustrating an example method for classifying an RMI flow in accordance with one embodiment of the present invention.

15 As shown in FIG. 5, processing begins in the client system 102 in S502 where requests for access to an RMI transport layer are made, for example, by application library software. If an RMI connection is being requested initiated (determined in S504), processing advances to S506, where the flow information associated with the RMI connection is detected. In S508, the detected flow information is transmitted upstream to the edge router 106 for 20 dynamic QoS classification, and the client system 102 processing returns to detecting a new RMI connection in S502.

Processing continues in the edge router 106 in S510, where the received RMI flow information is used to update a dynamic RMI flow list. Processing continues to S512 where all incoming packets bound for the outside network 108 are screened at the edge router 106. If a packet belongs to an RMI flow (determined in S512), it is marked with a DiffServ label 5 in accordance with the desired QoS for RMI flows in S514. All packets, whether marked with DiffServ labels or not, are then forwarded along to the outside network 108 in S516 and the process of screening packets at the edge router 106 continues at S510.

The methods and implementing apparatus of the present invention have been described in connection with the preferred embodiments as disclosed herein. Although exemplary embodiments of the present invention have been shown and described in detail herein, along with certain variants thereof, other varied embodiments which incorporate the teachings of the invention may easily be constructed by those skilled in the art.

For example, the above-described embodiment may be implemented in a number of ways, including the use of dedicated hardware, a combination of dedicated hardware and programmed special purpose processors, programmed general purpose processors or software, and the like.

Accordingly, the present invention is not intended to be limited to the specific form set forth herein, but on the contrary, it is intended to cover such alternatives, modifications, and equivalents, as can be reasonably included within the spirit and scope of the invention. In 20 other instances, well known structures are not shown in detail but can readily constructed by those skilled in the art.